Canonical forms for complex matrix congruence and *congruence

Roger A. Horn
Department of Mathematics, University of Utah
Salt Lake City, Utah 84103, rhorn@math.utah.edu

Vladimir V. Sergeichuk*
Institute of Mathematics, Tereshchenkivska 3
Kiev, Ukraine, sergeich@imath.kiev.ua

Abstract

Canonical forms for congruence and *congruence of square complex matrices were given by Horn and Sergeichuk in [Linear Algebra Appl. 389 (2004) 347–353], based on Sergeichuk's paper [Math. USSR, Izvestiya 31 (3) (1988) 481–501], which employed the theory of representations of quivers with involution. We use standard methods of matrix analysis to prove directly that these forms are canonical. Our proof provides explicit algorithms to compute all the blocks and parameters in the canonical forms. We use these forms to derive canonical pairs for simultaneous congruence of pairs of complex symmetric and skew-symmetric matrices as well as canonical forms for simultaneous *congruence of pairs of complex Hermitian matrices.

AMS classification: 15A21; 15A63

Keywords: Canonical forms, Congruence, *Congruence, Bilinear forms; Sesquilinear forms, Canonical pairs.

1 Introduction

Canonical matrices for congruence and *congruence over any field $\mathbb{F}$ of characteristic not 2 were established in [20, Theorem 3] up to classification of

This is the authors' version of a work that was published in Linear Algebra Appl. 416 (2006) 1010–1032.

*The research was started while this author was visiting the University of Utah supported by NSF grant DMS-0070503.

Hermitian forms over finite extensions of $\mathbb{F}$. Canonical forms for complex matrix congruence and *congruence are special cases of the canonical matrices in [20] since a classification of Hermitian forms over the complex field is known. Simpler versions of these canonical forms were given in [8], which relied on [20], and hence on the theory of representations of quivers with involution on which the latter is based.

In this paper, all matrices considered are complex. We use standard tools of matrix analysis to give a direct proof that the complex matrices given in [8] are canonical for congruence and *congruence.

Let A and B be square complex matrices of the same size. We say that A and B are *congruent* if there is a nonsingular S such that $S^TAS = B$; they are **congruent* if there is a nonsingular S such that $S^*AS = B$. We let $S^* := [\bar{s}_{ji}] = \bar{S}^T$ denote the complex conjugate transpose of $S = [s_{ij}]$ and write $S^{-T} := (S^{-1})^T$ and $S^{-*} := (S^{-1})^*$.

Define the n-by-n matrices

$$\Gamma_n = \begin{bmatrix} 0 & & & & & (-1)^{n+1} \\ & & & & \iddots & (-1)^n \\ & & & -1 & \iddots & \\ & & 1 & 1 & & \\ & -1 & -1 & & & \\ 1 & 1 & & & & 0 \end{bmatrix} \quad (\Gamma_1 = [1]), \tag{1}$$

$$\Delta_n = \begin{bmatrix} 0 & & & 1 \\ & & \iddots & i \\ & 1 & \iddots & \\ 1 & i & & 0 \end{bmatrix} \quad (\Delta_1 = [1]), \tag{2}$$

and the n-by-n Jordan block with eigenvalue λ

$$J_n(\lambda) = \begin{bmatrix} \lambda & 1 & & 0 \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix} \quad (J_1(\lambda) = [\lambda]).$$

The most important properties of these matrices for our purposes are that Γ_n is real and $\Gamma_n^{-T}\Gamma_n = \Gamma_n^{-*}\Gamma_n$ is similar to $J_n((-1)^{n+1})$; Δ_n is symmetric and $\Delta_n^{-*}\Delta_n = \bar{\Delta}_n^{-1}\Delta_n$ is similar to $J_n(1)$.

We also define the $2n$-by-$2n$ matrix

$$H_{2n}(\mu) = \begin{bmatrix} 0 & I_n \\ J_n(\mu) & 0 \end{bmatrix} \quad \left(H_2(\mu) = \begin{bmatrix} 0 & 1 \\ \mu & 0 \end{bmatrix}\right), \tag{3}$$

the *skew sum* of $J_n(\mu)$ and I_n. If $\mu \neq 0$, then $H_{2n}(\mu)^{-T} H_{2n}(\mu)$ is similar to $J_n(\mu) \oplus J_n(\mu^{-1})$ and $H_{2n}(\mu)^{-*} H_{2n}(\mu)$ is similar to $J_n(\mu) \oplus J_n(\bar{\mu}^{-1})$.

Sylvester's Inertia Theorem describes the *congruence canonical form of a complex Hermitian matrix. Our main goal is to give a direct proof of the following theorem, which generalizes Sylvester's theorem to all square complex matrices.

Theorem 1 ([8, Section 2])**.** *(a) Each square complex matrix is congruent to a direct sum, uniquely determined up to permutation of summands, of canonical matrices of the three types*

Type 0	$J_n(0)$	
Type I	Γ_n	(4)
Type II	$H_{2n}(\mu)$, $\quad 0 \neq \mu \neq (-1)^{n+1}$, μ *is determined up to replacement by* μ^{-1}	

*(b) Each square complex matrix is *congruent to a direct sum, uniquely determined up to permutation of summands, of canonical matrices of the three types*

Type 0	$J_n(0)$	
Type I	$\lambda\Delta_n$, $\quad \lvert\lambda\rvert = 1$	(5)
Type II	$H_{2n}(\mu)$, $\quad \lvert\mu\rvert > 1$	

Instead of Δ_n*, one may use* Γ_n *or any other nonsingular* $n \times n$ *matrix* F_n *for which there exists a real* θ_n *such that* $F_n^{-*} F_n$ *is similar to* $J_n(e^{i\theta_n})$*.*

For *congruence canonical matrices of Type I, it is sometimes convenient to identify the unit-modulus canonical parameter λ with the *ray* $\{t\lambda : t > 0\}$ or with the *angle* θ such that $\lambda = e^{i\theta}$ and $0 \leq \theta < 2\pi$. If λ occurs as a coefficient of exactly k blocks Δ_n in a *congruence canonical form, we say that it is a *canonical angle (or ray) of order* n *with multiplicity* k.

Our proof of Theorem 1 provides explicit algorithms to compute the sizes and multiplicities of the canonical blocks $J_n(0)$, Γ_n, $\lambda\Delta_n$, and $H_{2n}(\mu)$ and their canonical parameters λ and μ.

It suffices to prove Theorem 1 only for nonsingular matrices because of the following lemma, which is a specialization to the complex field of a *regularizing decomposition* for square matrices over any field or skew field with an involution [9].

Lemma 2. *Each square complex matrix A is congruent (respectively, *congruent) to a direct sum of the form*

$$B \oplus J_{r_1}(0) \oplus \cdots \oplus J_{r_p}(0) \quad \textit{with a nonsingular } B. \tag{6}$$

*This direct sum is uniquely determined up to permutation of its singular direct summands and replacement of B by any matrix that is congruent (respectively, *congruent) to it.*

The nonsingular direct summand B in (6) is called the *regular part* of A; the singular blocks in (6) are its Type 0 blocks. There is a simple algorithm to determine all of the direct summands in (6). If desired, this algorithm can be carried out using only unitary transformations [9].

In our development, it is convenient to use some basic properties of *primary matrix functions.* For a given square complex matrix A and a given complex valued function f that is analytic on a suitable open set containing the spectrum of A, the primary matrix function $f(A)$ may be defined using a power series, an explicit formula involving the Jordan canonical form of A, or a contour integral. For our purposes, its most important property is that for each A, $f(A)$ is a polynomial in A (the polynomial may depend on A, however), so $f(A)$ commutes with any matrix that commutes with A. For a systematic exposition of the theory of primary matrix functions, see [6, Chapter 6].

Preceded by Weierstrass, Kronecker developed a comprehensive theory of equivalence of matrix pencils in the late nineteenth century, but a similarly complete theory of matrix congruence has been achieved only recently. Gabriel [3] reduced the problem of equivalence of bilinear forms to the problem of equivalence of nonsingular bilinear forms. Riehm [16] reduced the problem of equivalence of nonsingular bilinear forms to the problem of equivalence of Hermitian forms. His reduction was improved and extended to sesquilinear forms in [17].

Using Riehm's reduction, Corbas and Williams [2] studied canonical forms for matrix congruence over an algebraically closed field with characteristic not 2. However, their proposed nonsingular canonical matrices are cumbersome and not canonical, e.g., their matrices

$$A = \begin{bmatrix} 0 & 1 \\ 1/2 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}$$

are actually congruent: $BAB^T = B$. For the singular case, they refer to the list of "singular blocks of known type" in [23, p. 60]. These singular

blocks are canonical but cumbersome, and we are fortunate that they may be replaced by the set of singular Jordan blocks; see [20] or [9].

Any square complex matrix A can be represented uniquely as $A = \mathcal{S} + \mathcal{C}$, in which $\mathcal{S}$ is symmetric and $\mathcal{C}$ is skew-symmetric; it can also be represented uniquely as $A = \mathcal{H} + i\mathcal{K}$, in which both $\mathcal{H}$ and $\mathcal{K}$ are Hermitian. A simultaneous congruence of $\mathcal{S}$ and $\mathcal{C}$ corresponds to a congruence of A and a simultaneous *congruence of $\mathcal{H}$ and $\mathcal{K}$ corresponds to a *congruence of A. Thus, if one has canonical forms for $\mathcal{S}$ and $\mathcal{C}$ under simultaneous congruence (often called *canonical pairs*), then one can obtain a canonical form for A under congruence as a consequence. Similarly, a canonical form for A under *congruence can be obtained if one has canonical pairs for two Hermitian matrices under simultaneous *congruence. Canonical pairs of both types may be found in Thompson's landmark paper [21] as well as in Lancaster and Rodman's recent reviews [14] and [13]. Thompson's canonical pairs were used to obtain canonical matrices for congruence over the real field by Lee and Weinberg [15], who observed that "the complex case follows from Thompson's results just as easily."

However, deriving canonical forms for complex congruence and *congruence from canonical pairs is like deriving the theory of conformal mappings in the real plane from properties of conjugate pairs of real harmonic functions. It can be done, but there are huge technical and conceptual advantages to working with complex analytic functions of a complex variable instead. Likewise, to derive congruence or *congruence canonical forms for a complex matrix A we advocate working directly with A rather than with its associated pairs $(\mathcal{S}, \mathcal{C})$ or $(\mathcal{H}, \mathcal{K})$. Our approach leads to three simple canonical block types for complex congruence rather than the six cumbersome block types found by Lee and Weinberg [15, p. 208].

Of course, canonical pairs for $(\mathcal{S}, \mathcal{C})$ and $(\mathcal{H}, \mathcal{K})$ follow from congruence and *congruence canonical forms for A. Define the following n-by-n matrices:

$$M_n := \begin{bmatrix} 0 & 1 & & 0 \\ 1 & 0 & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & 1 & 0 \end{bmatrix}, \quad N_n := \begin{bmatrix} 0 & 1 & & 0 \\ -1 & 0 & \ddots & \\ & \ddots & \ddots & 1 \\ 0 & & -1 & 0 \end{bmatrix},$$

$$X_n := \begin{bmatrix} 0 & & & & & (-1)^{n+1} \\ & & & & \cdot^{\cdot^{\cdot}} & 0 \\ & & & -1 & \cdot^{\cdot^{\cdot}} & \\ & & 1 & 0 & & \\ & -1 & 0 & & & \\ 1 & 0 & & & & 0 \end{bmatrix}, \quad Y_n := \begin{bmatrix} 0 & & & & & 0 \\ & & & & \cdot^{\cdot^{\cdot}} & (-1)^n \\ & & & 0 & \cdot^{\cdot^{\cdot}} & \\ & & 0 & 1 & & \\ & 0 & -1 & & & \\ 0 & 1 & & & & 0 \end{bmatrix},$$

and a 2-parameter version of the matrix (2)

$$\Delta_n(a,b) := \begin{bmatrix} 0 & & & a \\ & & \cdot^{\cdot^{\cdot}} & b \\ & a & \cdot^{\cdot^{\cdot}} & \\ a & b & & 0 \end{bmatrix}, \qquad a, b \in \mathbb{C}.$$

Define the *direct sum* of two matrix pairs

$$(A_1, A_2) \oplus (B_1, B_2) := (A_1 \oplus B_1,\ A_2 \oplus B_2)$$

and the *skew sum* of two matrices

$$[A \diagdown B] := \begin{bmatrix} 0 & B \\ A & 0 \end{bmatrix}.$$

Matrix pairs (A_1, A_2) and (B_1, B_2) are said to be *simultaneously congruent* (respectively, *simultaneously *congruent*) if there is a nonsingular matrix R such that $A_1 = R^T B_1 R$ and $A_2 = R^T B_2 R$ (respectively, $A_1 = R^* B_1 R$ and $A_2 = R^* B_2 R$). This transformation is a *simultaneous congruence* (respectively, *a simultaneous *congruence*) *of the pair* (B_1, B_2) *via* R.

The following theorem lists the canonical pairs that can occur and their associations with the congruence and *congruence canonical matrices listed in (4) and (5) of Theorem 1. The parameters λ and μ are as described in Theorem 1; the parameters ν, c, a, and b in the canonical pairs are functions of λ and μ.

Theorem 3. *(a) Each pair* $(\mathcal{S}, \mathcal{C})$ *consisting of a symmetric complex matrix* $\mathcal{S}$ *and a skew-symmetric complex matrix* $\mathcal{C}$ *of the same size is simultaneously congruent to a direct sum of pairs, determined uniquely up to permutation of summands, of the following three types, each associated with the indicated congruence canonical matrix type for* $A = \mathcal{S} + \mathcal{C}$*:*

Type 0: $J_n(0)$	$(M_n,\, N_n)$
Type I: Γ_n	(X_n, Y_n) *if* n *is odd,* (Y_n, X_n) *if* n *is even*
Type II: $H_{2n}(\mu)$	$([J_n(\mu+1)\diagdown J_n(\mu+1)^T],$ $[J_n(\mu-1)\diagdown -J_n(\mu-1)^T])$
$0 \neq \mu \neq (-1)^{n+1}$ *and* μ *is determined up to replacement by* μ^{-1}	

(7)

The Type II pair in (7) *can be replaced by two alternative pairs*

Type II: $H_{2n}(\mu)$ $0 \neq \mu \neq -1$ $\mu \neq 1$ *if* n *is odd*	$\big([I_n\diagdown I_n],\, [J_n(\nu)\diagdown -J_n(\nu)^T]\big)$ $\nu \neq 0$ *if* n *is odd,* $\nu \neq \pm 1,$ ν *is determined up to replacement by* $-\nu$
Type II: $H_{2n}(-1)$ n *is odd*	$\big([J_n(0)\diagdown J_n(0)^T],\, [I_n\diagdown -I_n]\big)$ n *is odd*

(8)

in which

$$\nu = \frac{\mu-1}{\mu+1}.$$

(b) Each pair $(\mathcal{H}, \mathcal{K})$ *of Hermitian matrices of the same size is simultaneously *congruent to a direct sum of pairs, determined uniquely up to permutation of summands, of the following four types, each associated with the indicated congruence canonical matrix type for* $A = \mathcal{H} + i\mathcal{K}$*:*

Type 0: $J_n(0)$	$(M_n,\, iN_n)$
Type I: $\lambda\Delta_n$ $\lvert\lambda\rvert = 1,\ \lambda^2 \neq -1$	$\pm\,(\Delta_n(1,0),\ \Delta_n(c,1))$ $c \in \mathbb{R}$
Type I: $\lambda\Delta_n$ $\lambda^2 = -1$	$\pm\,(\Delta_n(0,1),\ \Delta_n(1,0))$
Type II: $H_{2n}(\mu)$ $\lvert\mu\rvert > 1$	$([I_n \diagdown I_n],\ [J_n(a+ib)\diagdown J_n(a+ib)^*])$ $a, b \in \mathbb{R},\ a+bi \neq i,\ b > 0$

(9)

in which

$$c = \frac{\operatorname{Im}\lambda}{\operatorname{Re}\lambda}, \qquad a = \frac{2\operatorname{Im}\mu}{|1+\mu|^2}, \qquad b = \frac{|\mu|^2-1}{|1+\mu|^2}.$$

2 Congruence

The *cosquare* of a nonsingular complex matrix A is $A^{-T}A$. If B is congruent to A, then $A = S^T BS$ for some nonsingular S and hence

$$A^{-T}A = (S^T BS)^{-T}(S^T BS) = S^{-1}B^{-T}S^{-T}S^T BS = S^{-1}(B^{-T}B)S,$$

so *congruent nonsingular matrices have similar cosquares*. The following lemma establishes the converse assertion; an analogous statement for arbitrary systems of forms and linear mappings was given in [19] and [20, Theorem 1 and § 2].

Lemma 4. *Nonsingular complex matrices A and B are congruent if and only if their cosquares are similar.*

Proof. Let $A^{-T}A$ and $B^{-T}B$ be similar via S, so that

$$A^{-T}A = S^{-1}B^{-T}BS = \left(S^{-1}B^{-T}S^{-T}\right)\left(S^T BS\right) = C^{-T}C, \tag{10}$$

in which $C := S^T BS$. It suffices to prove that C is congruent to A. Let $M := CA^{-1}$ and deduce from (10) that

$$M = CA^{-1} = C^T A^{-T} = \left(A^{-1}C\right)^T \text{ and } M^T = A^{-1}C.$$

Thus,

$$C = MA = AM^T$$

and hence

$$q(M)A = Aq(M^T) = Aq(M)^T$$

for any polynomial $q(t)$. The theory of primary matrix functions [6, Section 6.4] ensures that there is a polynomial $p(t)$ such that $p(M)^2 = M$, so $p(M)$ is nonsingular and

$$p(M)A = Ap(M)^T.$$

Thus,

$$C = MA = p(M)^2 A = p(M)Ap(M)^T$$

so C is congruent to A via $p(M)$. □

Proof of Theorem 1(a). Let A be square and nonsingular. The Jordan Canonical Form of $A^{-T}A$ has a very special structure:

$$\bigoplus_{i=1}^{p}\left(J_{m_i}(\mu_i) \oplus J_{m_i}(\mu_i^{-1})\right) \oplus \bigoplus_{j=1}^{q} J_{n_j}((-1)^{n_j+1}), \quad 0 \neq \mu_i \neq (-1)^{m_i+1}; \tag{11}$$

see [22, Theorem 2.3.1] or [1, Theorem 3.6]. Using (11), form the matrix

$$B = \bigoplus_{i=1}^{p} H_{2m_i}(\mu_i) \oplus \bigoplus_{j=1}^{q} \Gamma_{n_j}.$$

Since the cosquare

$$H_{2m}(\mu)^{-T} H_{2m}(\mu) = \begin{bmatrix} 0 & I_m \\ J_m(\mu)^{-T} & 0 \end{bmatrix} \begin{bmatrix} 0 & I_m \\ J_m(\mu) & 0 \end{bmatrix} = \begin{bmatrix} J_m(\mu) & 0 \\ 0 & J_m(\mu)^{-T} \end{bmatrix}$$

is similar to $J_m(\mu) \oplus J_m(\mu^{-1})$ and the cosquare

$$\Gamma_n^{-T}\Gamma_n = (-1)^{n+1} \begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \iddots \\ -1 & -1 & -1 & -1 & \\ 1 & 1 & 1 & & \\ -1 & -1 & & & \\ 1 & & & & 0 \end{bmatrix}^T \cdot \Gamma_n = (-1)^{n+1} \begin{bmatrix} 1 & 2 & & \star \\ & 1 & \ddots & \\ & & \ddots & 2 \\ 0 & & & 1 \end{bmatrix}$$

is similar to $J_n((-1)^{n+1})$, (11) is the Jordan Canonical Form of $B^{-T}B$. Hence A and B have similar cosquares. Lemma 4 now ensures that A and B are congruent. Moreover, for any C that is congruent to A, its cosquare $C^{-T}C$ is similar to $A^{-T}A$, so it has the Jordan Canonical Form (11), which is uniquely determined up to permutation of summands. Hence B is uniquely determined up to permutation of its direct summands. □

Our proof of Theorem 1(a) shows that the congruence canonical form of a square complex matrix A can be constructed as follows:

1. Use the regularizing algorithm in [9, Section 2] to construct a regularizing decomposition (6) of A (if desired, one may use only unitary transformations in that algorithm).

2. Let B be the regular part of A and determine the Jordan canonical form (11) of its cosquare $B^{-T}B$.

3. Then
$$\bigoplus_{i=1}^{p} H_{2m_i}(\mu_i) \oplus \bigoplus_{j=1}^{q} \Gamma_{n_j} \oplus J_{r_1}(0) \oplus \cdots \oplus J_{r_p}(0)$$
is the congruence canonical form of A.

3 *Congruence

The **cosquare* of a nonsingular complex matrix A is $\mathcal{A} = A^{-*}A$. If B is *congruent to A, then $A = S^*BS$ for some nonsingular S and hence

$$A^{-*}A = (S^*BS)^{-*}(S^*BS) = S^{-1}B^{-*}S^{-*}S^*BS = S^{-1}(B^{-*}B)S,$$

so *congruent nonsingular matrices have similar *cosquares. However, $[-1]$ is the *cosquare of both $[i]$ and $[-i]$, which are not *congruent: there is no nonzero complex s such that $-i = \bar{s}is = |s|^2 i$. Nevertheless, there is a useful analog of Lemma 4 for *congruence. We denote the set of *distinct* eigenvalues of a square matrix X by $\operatorname{dspec} X$.

Lemma 5. *Let A and B be nonsingular n-by-n complex matrices with similar *cosquares, that is, $A^{-*}A = S^{-1}(B^{-*}B)S$ for some nonsingular S. Let $B_S := S^*BS$, let $M := B_S A^{-1}$, and suppose M has k real negative eigenvalues, counted according to their algebraic multiplicities $(0 \le k \le n)$. Then:*
(a) M is similar to a real matrix.
*(b) There are square complex matrices D_- and D_+ of size k and $n-k$, respectively, such that A is *congruent to $(-D_-) \oplus D_+$ and B is *congruent to $D_- \oplus D_+$.*

Proof. We have

$$A^{-*}A = S^{-1}(B^{-*}B)S = \left(S^{-1}B^{-*}S^{-*}\right)(S^*BS) = B_S^{-*}B_S,$$

from which it follows that

$$M = B_S A^{-1} = B_S^* A^{-*} = \left(A^{-1}B_S\right)^* \quad \text{and} \quad M^* = A^{-1}B_S,$$

and hence

$$B_S = MA = AM^*. \tag{12}$$

Thus, $M^* = A^{-1}MA$, so M is similar to a real matrix [5, Theorem 4.1.7]. Its Jordan blocks with nonreal eigenvalues occur in conjugate pairs, so there is a nonsingular T such that $TMT^{-1} = M_- \oplus M_+$, in which the k-by-k matrix M_- is either absent or has only real negative eigenvalues; M_+ has no negative eigenvalues and is similar to a real matrix if it is present. Moreover, if we partition $TAT^* = [A_{ij}]_{i,j=1}^2$ conformally to $M_- \oplus M_+$, Sylvester's Theorem on linear matrix equations [6, Theorem 4.4.6] ensures that

$$TAT^* = A_{11} \oplus A_{22} \tag{13}$$

since the equalities $\operatorname{dspec} M_- \cap \operatorname{dspec} M_+ = \varnothing$, $\operatorname{dspec} M_- = \operatorname{dspec} M_-^*$, and $\operatorname{dspec} M_+ = \operatorname{dspec} M_+^*$ imply that

$$\begin{aligned} TB_ST^* &= \left(TMT^{-1}\right)(TAT^*) = (TAT^*)\left(T^{-*}M^*T^*\right) \\ &= (M_- \oplus M_+)(TAT^*) = (TAT^*)\left(M_-^* \oplus M_+^*\right) \\ &= \begin{bmatrix} M_-A_{11} & M_-A_{12} \\ M_+A_{21} & M_+A_{22} \end{bmatrix} = \begin{bmatrix} A_{11}M_-^* & A_{12}M_+^* \\ A_{21}M_-^* & A_{22}M_+^* \end{bmatrix} \\ &= M_-A_{11} \oplus M_+A_{22} = A_{11}M_-^* \oplus A_{22}M_+^*. \end{aligned}$$

Thus,

$$(-M_-)]A_{11} = A_{11}(-M_-^*) \quad \text{and} \quad M_+A_{22} = A_{22}M_+^*,$$

so

$$q_1(-M_-)A_{11} = A_{11}q_1(-M_-^*) = A_{11}q_1(-M_-)^*$$

and

$$q_2(M_+)A_{22} = A_{22}q_2(M_+^*) = A_{22}q_2(M_+)^*$$

for any polynomials $q_1(t)$ and $q_2(t)$ with real coefficients. Neither $-M_-$ nor M_+ has any negative eigenvalues and each is similar to a real matrix, so [7, Theorem 2(c)] ensures that there are polynomials $g(t)$ and $h(t)$ with real coefficients such that $g(-M_-)^2 = -M_-$ and $h(M_+)^2 = M_+$. It follows that B is *congruent to

$$\begin{aligned} TB_ST^* &= -(-M_-A_{11}) \oplus M_+A_{22} = -g(M_-)^2A_{11} \oplus h(M_+)^2A_{22} \\ &= -g(-M_+)A_{11}g(-M_+)^* \oplus h(M_+)A_{22}h(M_+)^* = D_- \oplus D_+, \end{aligned}$$

in which $D_- = -g(-M_+)A_{11}g(-M_+)^*$ and $D_+ = h(M_+)A_{22}h(M_+)^*$; A is *congruent to $(-D_-) \oplus D_+$ by (13) and B is *congruent to $D_- \oplus D_+$. $\square$

The result cited from [7, Theorem 2(c)] gives sufficient conditions for $f(X) = Y$ to have a real solution X for a given real Y. The key conditions are that f is analytic and one-to-one on a domain that is symmetric with respect to the real axis and f^{-1} is typically real, that is, $f(\bar{z}) = \overline{f(z)}$ on the range of f. Under these conditions there is a solution X that is a polynomial in Y with real coefficients. In the case at hand, $f(z) = z^2$ on the open right half-plane; this special case appears in [10, p. 545] and was employed in [21, p. 356] and [14, Lemma 7.2] to study canonical pairs of Hermitian matrices.

Two noteworthy special cases of Lemma 5 occur when M either has only positive eigenvalues ($k = n$) or only negative eigenvalues ($k = 0$). In the former case, A is *congruent to B; in the latter case, A is *congruent to $-B$.

Proof of Theorem 1(b): Existence. Let A be nonsingular and let $\mathcal{A} = A^{-*}A$ denote its *cosquare. Because $\mathcal{A}^{-*} = (A^{-*}A)^{-*} = AA^{-*} = A\mathcal{A}A^{-1}$, for each eigenvalue λ of $\mathcal{A}$ and each $k = 1, 2, \ldots$, $J_k(\lambda)$ and $J_k(\bar{\lambda}^{-1})$ have equal multiplicities in the Jordan Canonical Form of $\mathcal{A}$. Since $\lambda = \bar{\lambda}^{-1}$ whenever $|\lambda| = 1$, this pairing is trivial for any eigenvalue of $\mathcal{A}$ that has modulus one; it is nontrivial for eigenvalues $\mathcal{A}$ whose modulus is greater than one.

Let

$$\bigoplus_{i=1}^{p} \left(J_{m_i}(\mu_i) \oplus J_{m_i}(\bar{\mu}_i^{-1})\right) \oplus \bigoplus_{j=1}^{q} J_{n_j}(e^{i\phi_j}), \quad |\mu_i| > 1,\ 0 \le \phi_j < 2\pi \tag{14}$$

be the Jordan Canonical Form of $\mathcal{A}$ and use it to construct the matrix

$$B := \bigoplus_{i=1}^{p} H_{2m_i}(\mu_i) \oplus \bigoplus_{j=1}^{q} e^{i\phi_j/2}\Delta_{n_j}. \tag{15}$$

Since the *cosquare

$$H_{2m}(\mu)^{-*}H_{2m}(\mu) = \begin{bmatrix} 0 & I_m \\ J_m(\mu)^{-*} & 0 \end{bmatrix} \begin{bmatrix} 0 & I_m \\ J_m(\mu) & 0 \end{bmatrix} = \begin{bmatrix} J_m(\mu) & 0 \\ 0 & J_m(\mu)^{-*} \end{bmatrix}$$

is similar to $J_m(\mu) \oplus J_m(\bar{\mu}^{-1})$ and the *cosquare

$$\Delta_n^{-*}\Delta_n = \begin{bmatrix} 1 & 2i & & \star \\ & 1 & \ddots & \\ & & \ddots & 2i \\ 0 & & & 1 \end{bmatrix} \tag{16}$$

is similar to $J_n(1)$, (14) is also the Jordan Canonical Form of $B^{-*}B$. Hence A and B have similar *cosquares, whose common Jordan Canonical Form contains $2p + q$ Jordan blocks (p block pairs of the form $J_m(\mu) \oplus J_m(\bar{\mu}^{-1})$ with $|\mu| > 1$ and q blocks of the form $\lambda\Delta_n$ with $|\lambda| = 1$).

Lemma 5 ensures that there are matrices D_+ and D_- such that A is *congruent to $(-D_-) \oplus D_+$ and B is *congruent to $D_- \oplus D_+$. If D_- is absent, then A is *congruent to B; if D_+ is absent, then A is *congruent to $-B$. In both of these cases A is *congruent to a direct sum of the form

$$\bigoplus_{i=1}^{p} \varepsilon_i H_{2m_i}(\mu_i) \oplus \bigoplus_{j=1}^{q} \delta_j e^{i\phi_j/2}\Delta_{n_j}, \quad \varepsilon_i, \delta_j \in \{-1, +1\}. \tag{17}$$

If both direct summands D_- and D_+ are present, then their sizes are less than the size of A. Reasoning by induction, we may assume that each of

D_- and D_+ is *congruent to a direct sum of the form (17). Then A is *congruent to a direct sum of the form (17) as well. We may take all $\varepsilon_i = 1$ in (17) since each $H_{2m}(\mu)$ is *congruent to $-H_{2m}(\mu)$:

$$\begin{bmatrix} I_m & 0 \\ 0 & -I_m \end{bmatrix} \begin{bmatrix} 0 & I_m \\ J_m(\mu) & 0 \end{bmatrix} \begin{bmatrix} I_m & 0 \\ 0 & -I_m \end{bmatrix} = - \begin{bmatrix} 0 & I_m \\ J_m(\mu) & 0 \end{bmatrix}.$$

□

We have demonstrated that a nonsingular A is *congruent to a direct sum of Type I and Type II blocks

$$\bigoplus_{i=1}^{p} H_{2m_i}(\mu_i) \oplus \bigoplus_{j=1}^{q} \delta_j e^{i\phi_j/2} \Delta_{n_j}, \quad \delta_j \in \{-1, +1\},\ |\mu_i| > 1,\ 0 \le \phi_j < 2\pi, \tag{18}$$

in which the sizes $2m_i$ and parameters μ_i of the Type II blocks as well as the sizes n_j and squared parameters $(\delta_j e^{i\phi_j/2})^2 = e^{i\phi_j}$ of the Type I blocks are uniquely determined by A. Our reduction algorithm using Lemma 5 determines a set of signs $\{\delta_j\}$ that gives the desired *congruence of A to (18), but we must show that no other choice of signs is possible: in the set of Type I blocks in (18) with equal sizes n_j and equal coefficients $e^{i\phi_j/2}$ the number of blocks with sign equal to $+1$ (and hence also the number of blocks with signs equal to -1) is uniquely determined by A.

Proof of Theorem 1(b): Uniqueness. Let each of A and B be a direct sum of Type I and Type II blocks and suppose that A and B are *congruent. We have shown that A and B have the form

$$A = \bigoplus_{i=1}^{p} H_{2m_i}(\mu_i) \oplus \bigoplus_{j=1}^{q} \lambda_j \Delta_{n_j}, \quad B = \bigoplus_{i=1}^{p} H_{2m_i}(\mu_i) \oplus \bigoplus_{j=1}^{q} \kappa_j \lambda_j \Delta_{n_j},$$

in which all $\kappa_j \in \{-1, +1\}$, all $|\mu_i| > 1$, and all $|\lambda_j| = 1$. Our goal is to prove that each of these direct sums may be obtained from the other by a permutation of summands. We may rearrange the summands to present $A = A_1 \oplus A_2$ and $B = B_1 \oplus B_2$, in which

$$A_1 = \bigoplus_{r=1}^{k} \lambda_r \Delta_{n_r}, \quad B_1 = \bigoplus_{r=1}^{k} \kappa_r \lambda_r \Delta_{n_r},$$

and $\lambda_1^2 = \cdots = \lambda_k^2 \neq \lambda_\ell^2$ for all $\ell = k+1, \ldots, q$. Let $S^*AS = B$ and partition $S = [S_{ij}]_{i,j=1}^2$ conformally with $A_1 \oplus A_2$. Since the *cosquares of A and B are similar via S, we have $S\left(B^{-*}B\right) = \left(A^{-*}A\right)S$ and hence

$$\begin{bmatrix} S_{11}\left(B_1^{-*}B_1\right) & S_{12}\left(B_2^{-*}B_2\right) \\ S_{21}\left(B_1^{-*}B_1\right) & S_{22}\left(B_2^{-*}B_2\right) \end{bmatrix} = \begin{bmatrix} \left(A_1^{-*}A_1\right)S_{11} & \left(A_1^{-*}A_1\right)S_{12} \\ \left(A_2^{-*}A_2\right)S_{21} & \left(A_2^{-*}A_2\right)S_{22} \end{bmatrix}.$$

But

$$\operatorname{dspec}\left(B_2^{-*}B_2\right) \cap \operatorname{dspec}\left(A_1^{-*}A_1\right) = \varnothing$$

and

$$\operatorname{dspec}\left(B_1^{-*}B_1\right) \cap \operatorname{dspec}\left(A_2^{-*}A_2\right) = \varnothing,$$

so Sylvester's Theorem on linear matrix equations ensures that $S = S_{11} \oplus S_{22}$. Thus, A_1 is *congruent to B_1 via S_{11} and hence it suffices to consider the case $A = A_1$ and $B = B_1$. Moreover, dividing both A and B by λ_1 it suffices to consider a pair of *congruent matrices of the form

$$A = \bigoplus_{r=1}^{k} \varepsilon_r \Delta_{n_r},\ B = \bigoplus_{r=1}^{k} \delta_r \Delta_{n_r}, \quad \varepsilon_r, \delta_r \in \{-1, +1\}. \tag{19}$$

We may assume that the summands in (19) are arranged so that $1 \leq n_1 \leq \cdots \leq n_k$. Define

$$N := (n_1, \ldots, n_k) \quad \text{and} \quad |N| := n_1 + \cdots + n_k. \tag{20}$$

Let

$$J_N := J_{n_1}(0) \oplus \cdots \oplus J_{n_k}(0), \tag{21}$$

denote

$$\Delta_N := \Delta_{n_1} \oplus \cdots \oplus \Delta_{n_k} \quad \text{and} \quad \mathcal{D}_N := \Delta_N^{-*}\Delta_N,$$

and let

$$P_N := P_{n_1} \oplus \cdots \oplus P_{n_k},$$

in which

$$P_n := \begin{bmatrix} 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix}$$

is the n-by-n *reversal matrix*.

Since $P_N\Delta_N = I + iJ_N$ and J_N is nilpotent,

$$\begin{aligned} \mathcal{D}_N &= \Delta_N^{-*}\Delta_N = \left(P_N\Delta_N^*\right)^{-1} P_N\Delta_N = (I - iJ_N)^{-1}(I + iJ_N) \\ &= (I + iJ_N + i^2J_N^2 + i^3J_N^3 + \cdots)(I + iJ_N) \\ &= I + 2iJ_N + 2i^2J_N^2 + 2i^3J_N^3 + \cdots \end{aligned}$$

is a polynomial in J_N. Moreover, $J_N = i(I-\mathcal{D}_N)(I+\mathcal{D}_N)^{-1}$ and $(I+\mathcal{D}_N)^{-1}$ is a polynomial in $\mathcal{D}_N$, so J_N is a polynomial in $\mathcal{D}_N$.

Let C be $|N|$-by-$|N|$ and partition $C = [C_{ij}]_{i,j=1}^k$ conformally to J_N, so each block C_{ij} is n_i-by-n_j. Then C commutes with J_N if and only if each block C_{ij} has the form

$$C_{ij} = \begin{cases} \begin{bmatrix} & c_{ij} & c_{ij}^{(2)} & \dots & c_{ij}^{(n_i)} \\ & & c_{ij} & \ddots & \vdots \\ & & & \ddots & c_{ij}^{(2)} \\ 0 & & & & c_{ij} \end{bmatrix} & \text{if } i \le j, \\ \begin{bmatrix} c_{ij} & c_{ij}^{(2)} & \dots & c_{ij}^{(n_j)} \\ & c_{ij} & \ddots & \vdots \\ & & \ddots & c_{ij}^{(2)} \\ & & & c_{ij} \\ & & & \\ 0 & & & \end{bmatrix} & \text{if } i > j; \end{cases} \tag{22}$$

see [4, Section VIII, § 2] or [6, Lemma 4.4.11]. Each diagonal block C_{ii} is upper Toeplitz; each block C_{ij} with $i < j$ has an upper Toeplitz submatrix that is preceded by a zero block; each block C_{ij} with $i > j$ has an upper Toeplitz submatrix with a zero block below it. An $|N|$-by-$|N|$ matrix whose blocks have the form (22) is said to be *N-upper Toeplitz*.

If C is N-upper Toeplitz, then so is $P_N C^* P_N$; the upper Toeplitz submatrix in its i,j block is the complex conjugate of the upper Toeplitz submatrix in the j,i block of C. The matrices $\mathcal{D}_N$, $P_N A$, and $P_N B$ (see (19)) are all N-upper Toeplitz.

Let $C = [C_{ij}]_{i,j=1}^k be$ a given N-upper Toeplitz matrix. Consider the mapping $C \mapsto \underline{C}$ that takes C into the $k \times k$ matrix whose i,j entry is c_{ij} if $n_i = n_j$ and is 0 otherwise; see (22). Partition the entries of N (already nondecreasingly ordered) into groups of equal entries

$$1 \le n_1 = \cdots = n_r < n_{r+1} = \cdots = n_l < n_{l+1} = \cdots = n_t < \cdots$$

and observe that $\underline{C}$ is structurally block diagonal:

$$\underline{C} = C_1 \oplus C_2 \oplus C_3 \oplus \cdots \tag{23}$$

The sizes of the direct summands of $\underline{C}$ are the multiplicities of the entries of N, that is,

$$C_1 := \begin{bmatrix} c_{11} & \dots & c_{1r} \\ \vdots & \ddots & \vdots \\ c_{r1} & \dots & c_{rr} \end{bmatrix}, \quad C_2 := \begin{bmatrix} c_{r+1,r+1} & \dots & c_{r+1,\ell} \\ \vdots & \ddots & \vdots \\ c_{\ell,r+1} & \dots & c_{\ell\ell} \end{bmatrix}, \dots \tag{24}$$

In addition, for any N-upper Toeplitz matrix D we have $\underline{CD} = \underline{C} \cdot \underline{D}$. If C is nonsingular, we have $I = \underline{CC^{-1}} = \underline{C} \cdot \underline{C^{-1}}$, which implies that $\underline{C}$ is nonsingular. A computation reveals that $\underline{P_N C^* P_N} = (\underline{C})^*$.

Finally, consider the *congruent matrices A and B in (19), which satisfy $A^{-*}A = B^{-*}B = \mathcal{D}_N$. We have $\underline{P_N A} = \operatorname{diag}(\varepsilon_1, \dots, \varepsilon_k)$ and $\underline{P_N B} = \operatorname{diag}(\delta_1, \dots, \delta_k)$. Let $S^*AS = B$, so $B^{-*}B = S^{-1}(A^{-*}A)S$ and hence

$$S(B^{-*}B) = S\mathcal{D}_N = \mathcal{D}_N S = \left(A^{-*}A\right) S.$$

Since S commutes with $\mathcal{D}_N$ and J_N is a polynomial in $\mathcal{D}_N$, S commutes with J_N and hence is N-upper Toeplitz. Moreover,

$$P_N B = P_N S^* A S = (P_N S^* P_N)(P_N A) S$$

and each of the matrices $P_N B$, $P_N A$, S, and $P_N S^* P_N$ is N-upper Toeplitz. Therefore,

$$\underline{P_N B} = \left(\underline{P_N S^* P_N}\right)\left(\underline{P_N A}\right)\underline{S},$$

that is,

$$\operatorname{diag}(\delta_1, \dots, \delta_k) = \underline{S}^* \operatorname{diag}(\varepsilon_1, \dots, \varepsilon_k)\underline{S}.$$

But $\underline{S}$ is nonsingular and block diagonal:

$$\underline{S} = S_1 \oplus S_2 \oplus S_3 \oplus \cdots,$$

in which the respective nonsingular blocks $S_1, S_2, ...$ are the same size as the respective blocks $C_1, C_2, ...$ in (24). Thus, $\operatorname{diag}(\delta_1, \dots, \delta_r) = S_1^* \operatorname{diag}(\varepsilon_1, \dots, \varepsilon_r) S_1$, $\operatorname{diag}(\delta_{r+1}, \dots, \delta_\ell) = S_2^* \operatorname{diag}(\varepsilon_{r+1}, \dots, \varepsilon_\ell) S_2$, etc. Sylvester's Inertia Theorem ensures that $\operatorname{diag}(\delta_1, \dots, \delta_r)$ can be obtained from $\operatorname{diag}(\varepsilon_1, \dots, \varepsilon_r)$ by a permutation of its diagonal entries, $\operatorname{diag}(\delta_{r+1}, \dots, \delta_\ell)$ can be obtained from $\operatorname{diag}(\varepsilon_{r+1}, \dots, \varepsilon_\ell)$ by a permutation of its diagonal entries, etc. Therefore, each of the direct sums (19) can be obtained from the other by a permutation of summands. □

The argument that we have just made also clarifies the final assertion in Theorem 1: A *cosquare $F_n^{-*}F_n$ is similar to $J_n(\lambda)$ with $|\lambda| = 1$ if and only if it is not decomposable into a nontrivial direct sum under similarity.

4 An alternative algorithm for *congruence

Although we can now determine the *congruence canonical form of a nonsingular complex matrix A, in practice it is useful to have an alternative algorithm.

Let $\mathcal{A} = A^{-*}A$, let $\mu_1, \ldots, \mu_r$ be the distinct eigenvalues of $\mathcal{A}$ with modulus greater than one, and let $\lambda_1, \ldots, \lambda_s$ be the distinct eigenvalues of $\mathcal{A}$ with modulus one. Let S be any nonsingular matrix such that

$$A^{-*}A = S(C_1 \oplus \cdots \oplus C_r \oplus C_{r+1} \oplus \cdots \oplus C_{r+s})S^{-1}, \tag{25}$$

in which $\operatorname{dspec} C_i = \{\mu_i, \bar{\mu}_i^{-1}\}$ for $i = 1, \ldots, r$ and $\operatorname{dspec} C_{r+i} = \{\lambda_i\}$ for $i = 1, \ldots, s$. One way to achieve this decomposition is to group together blocks from the Jordan Canonical Form of $\mathcal{A}$, but other strategies may be employed. Partition $S^*AS = [A_{ij}]_{i,j=1}^{r+s}$ conformally to the direct sum in (25). The argument in the proof of uniqueness in Section 3 shows that S^*AS is block diagonal:

$$S^*AS = A_{11} \oplus \cdots \oplus A_{rr} \oplus A_{r+1,r+1} \oplus \cdots \oplus A_{s+1,s+1}, \tag{26}$$

in which each A_{ii} is the same size as C_i, $\operatorname{dspec} A_{ii}^{-*}A_{ii} = \{\mu_i, \bar{\mu}_i^{-1}\}$ for $i = 1, \ldots, r$, and $\operatorname{dspec} A_{ii}^{-*}A_{ii} = \{\lambda_{i-r}\}$ for $i = r+1, \ldots, r+s$.

The Type II *congruence blocks are now easy to determine: to each pair of Jordan blocks $J_m(\mu_i) \oplus J_m(\bar{\mu}_i^{-1})$ of $A_{ii}^{-*}A_{ii}$ corresponds one Type II *congruence block $H_{2m}(\mu_i)$ of A.

Now consider each diagonal block $A_{r+j,r+j}$ in turn. Its *cosquare has a single eigenvalue $\lambda_j = e^{i\phi_j}$, $0 \le \phi_j < 2\pi$. Let the Jordan Canonical Form of the *cosquare of $e^{-i\phi_j/2}A_{r+j,r+j}$ be $I + J_N$, in which $N := (n_1, \ldots, n_k)$ and $1 \le n_1 \le \cdots \le n_k$. That *cosquare is similar to $\mathcal{D}_N := \Delta_N^{-*}\Delta_N$; let S be nonsingular and such that $S\mathcal{D}_N S^{-1} = e^{-i\phi_j}A_{j+r,j+r}^{-*}A_{j+r,j+r}$. For notational convenience, normalize and set $A := e^{-i\phi_j/2}S^*A_{j+r,j+r}S$. Then

$$A^{-*}A = e^{-i\phi_j}S^{-1}A_{j+r,j+r}^{-*}A_{j+r,j+r}S = \mathcal{D}_N = \Delta_N^{-*}\Delta_N,$$

which implies that $A = A^*\Delta_N^{-*}\Delta_N$ and hence

$$\begin{aligned} B := \Delta_N^{-1}A &= \Delta_N^{-1}(A^*\Delta_N^{-*}\Delta_N) = \Delta_N^{-1}(\Delta_N^{-1}A)^*\Delta_N \\ &= \Delta_N^{-1}B^*\Delta_N = \Delta_N^{-1}\left(\Delta_N^*B\Delta_N^{-*}\right)\Delta_N = \mathcal{D}_N^{-1}B\mathcal{D}_N. \end{aligned} \tag{27}$$

Thus, $\mathcal{D}_N$ commutes with B, so B is N-upper Toeplitz.

Invoking the identity

$$B^* = \Delta_N B \Delta_N^{-1},$$

already employed in (27), compute

$$P_N B^* P_N = P_N \left(\Delta_N B \Delta_N^{-1}\right) P_N = \left(P_N \Delta_N\right) B \Delta_N^{-1} P_N.$$

Since $P_N \Delta_N = I + iJ_N$ commutes with any N-upper Toeplitz matrix, it follows that

$$P_N B^* P_N = B \left(P_N \Delta_N\right) \Delta_N^{-1} P_N = B P_N^2 = B. \tag{28}$$

Let $X = [X_{ij}]$ be a given N-upper Toeplitz matrix, partitioned as in (22). The *N-block star* of X is the N-upper Toeplitz matrix

$$X^{\boxast} := P_N X^* P_N.$$

The upper Toeplitz submatrix of each i, j block of $X^{\boxast}$ is the complex conjugate of the upper Toeplitz submatrix of X_{ji}. We say that X is *N-Hermitian* if $X^{\boxast} = X$, in which case the upper Toeplitz submatrices of each pair of blocks X_{ij} and X_{ji} are complex conjugates. If $N = (1, 1, \ldots, 1)$ then $X^{\boxast} = X^*$ and X is N-Hermitian if and only if it is Hermitian.

The identity (28) asserts that B is N-Hermitian.

If X and Y are N-upper Toeplitz, one checks that

$$(XY)^{\boxast} = Y^{\boxast} X^{\boxast}.$$

We say that N-upper Toeplitz matrices X and Y are ${}^{\boxast}$*congruent* (*N-block star congruent*) if there exists a nonsingular N-upper Toeplitz matrix S such that $S^{\boxast} B S = C$; ${}^{\boxast}$congruence is an equivalence relation on the set of N-upper Toeplitz matrices.

Since B is N-upper Toeplitz and N-Hermitian, for any N-upper Toeplitz matrix S we have

$$\begin{aligned} S^* A S = S^* \left(\Delta_N B\right) S &= \left(P_N S^{\boxast} P_N\right) \Delta_N B S = P_N S^{\boxast} (P_N \Delta_N) B S \\ &= P_N \left(P_N \Delta_N\right) S^{\boxast} B S = \Delta_N \left(S^{\boxast} B S\right). \end{aligned} \tag{29}$$

If we can find a nonsingular N-upper Toeplitz S such that

$$S^{\boxast} B S = \varepsilon_1 I_{n_1} \oplus \cdots \oplus \varepsilon_k I_{n_k} \quad \text{with } \varepsilon_i = \pm 1,$$

it follows from (29) that A is *congruent to

$$\varepsilon_1 \Delta_{n_1} \oplus \cdots \oplus \varepsilon_k \Delta_{n_k}. \tag{30}$$

Theorem 1(b) ensures that for each $n = 1, 2, \ldots$ there is a *unique* set of signs associated with the blocks Δ_n of size n in (30). The following generalization of Sylvester's Inertia Theorem provides a way to construct these signs.

Lemma 6. *Let C be nonsingular, N-upper Toeplitz, and N-Hermitian. Then there is a nonsingular N-upper Toeplitz matrix S such that*

$$S^{\boxed{*}}CS = \varepsilon_1 I_{n_1} \oplus \cdots \oplus \varepsilon_k I_{n_k}, \qquad \text{each } \varepsilon_i \in \{-1, 1\}. \tag{31}$$

Proof. Since C is nonsingular, $\underline{C}$ and hence all of the direct summands in (23) are nonsingular as well.

Step 1. If $c_{11} \neq 0$, proceed to Step 2. If $c_{11} = 0$, then $c_{1j} \neq 0$ for some $j \leq r$ since C_1 is nonsingular. Let $S_\theta = [S_{ij}]_{i,j=1}^k$ be the N-upper Toeplitz matrix in which the diagonal blocks are identity matrices and all the other blocks are zero except for $S_{j1} := e^{i\theta} I_{n_1}$ for a real θ to be determined. The $1,1$ entry of $S_\theta^{\boxed{*}} C S_\theta$ is

$$e^{i\theta} c_{1j} + e^{-i\theta} \bar{c}_{1j} + e^{i\theta} e^{-i\theta} c_{jj} = 2\,\mathrm{Re}(e^{i\theta} c_{1j}) + c_{jj} \tag{32}$$

($c_{j1} = \bar{c}_{1j}$ and $c_{jj} = \bar{c}_{jj}$ since $C^{\boxed{*}} = C$). Choose any θ for which (32) is nonzero.

Step 2. We may now assume that c_{11} is a nonzero real number. Let $a := |c_{11}|^{-1/2}$, so $a^2 c_{11} = \pm 1$. Define the real N-upper Toeplitz matrix $S = aI$ and form $S^{\boxed{*}}CS = SCS$, whose $1,1$ entry is ± 1.

Step 3. We may now assume that $\operatorname{dspec} C_{11} = \{c_{11}\} = \{\pm 1\}$. Then $\operatorname{dspec}(c_{11} C_{11}^{-1}) = \{1\}$, so there is a polynomial $p(t)$ with real coefficients such that $p(C_{11})^2 = c_{11} C_{11}^{-1}$ [6, Theorem 6.4.14], $p(C_{11})$ is upper Toeplitz and commutes with C_{11}, and $p(C_{11}) C_{11} p(C_{11}) = c_{11} I = \pm I_{n_1}$. Define the real N-upper Toeplitz matrix $S = p(C_{11}) \oplus I_{n_2} \oplus \cdots \oplus I_{n_k}$ and form $S^{\boxed{*}}CS$, whose $1,1$ block is $\pm I_{n_1}$.

Step 4. We may now assume that $C_{11} = \pm I_{n_1}$. Define the N-upper Toeplitz matrix

$$S = \begin{bmatrix} I_{n_1} & -c_{11} C_{12} & \dots & -c_{11} C_{1n_k} \\ & I_{n_2} & \dots & 0 \\ & & \ddots & \vdots \\ 0 & & & I_{n_k} \end{bmatrix}$$

and form $S^{\boxed{*}}CS$; its $1,j$ and $j,1$ blocks are zero for all $j = 2, \dots, n_k$.

The preceding four steps reduce C by $\boxed{*}$congruence to the form $\pm I_{n_1} \oplus C'$. Now reduce C' in the same way and continue. After k iterations of this process we obtain a real diagonal matrix

$$\varepsilon_1 I_{n_1} \oplus \cdots \oplus \varepsilon_k I_{n_k}, \quad \varepsilon_i \in \{-1, 1\} \tag{33}$$

that is $\boxed{*}$congruent to the original matrix C. □

Thus, to determine the *congruence canonical form of a given square complex matrix A, one may proceed as follows:

1. Apply the regularization algorithm [9] to determine the singular *congruence blocks and the regular part. This reduces the problem to consideration of a nonsingular A.

2. Let S be any nonsingular matrix that gives a similarity between the *cosquare of A and the direct sum in (25) and calculate S^*AS, which has the block diagonal form (26). Consider each of these diagonal blocks in turn.

3. Determine the Type II *congruence blocks by examining the Jordan Canonical Forms of the diagonal blocks whose *cosquare has a two-point spectrum: one Type II block $H_{2m}(\mu_i)$ corresponds to each pair $J_m(\mu_i) \oplus J_m(\bar{\mu}_i^{-1})$ in the Jordan Canonical Form of $A_{ii}^{-*}A_{ii}$.

4. For each diagonal block A_{jj} whose *cosquare has a one-point spectrum (suppose it is $e^{i\phi}$), consider $\hat{A}_{jj} = e^{-i\phi/2}A_{jj}$. Find an S such that $\hat{A}_{jj}^{-*}\hat{A}_{jj} = S^{-1}(\Delta_N^{-1}\Delta_N)S$ and consider $B = \Delta_N^{-1}S^*\hat{A}_{jj}S$, which is N-upper Toeplitz and N-Hermitian.

5. Use Lemma 6 (or some other means) to reduce B by *congruence to a diagonal form (33). Then $\hat{A}_{jj}$ is *congruent to a direct sum of the form (30) and the diagonal block A_{jj} corresponds to a direct sum $\varepsilon_1 e^{i\phi/2}\Delta_{n_1} \oplus \cdots \oplus \varepsilon_k e^{i\phi/2}\Delta_{n_k}$ of Type I blocks.

5 Some special *congruences

A square complex matrix A is diagonalizable by *congruence if and only if the *congruence canonical form of A contains only 1-by-1 blocks (which can only be Type 0 blocks $J_1(0) = [0]$ and Type I blocks $\lambda\Delta_1 = [\lambda]$ with $|\lambda| = 1$). The Type 0 *congruence canonical blocks for A are all 1-by-1 if and only if A and A^* have the same null space; an equivalent condition is that there is a unitary U such that

$$A = U^*(B \oplus 0_k)U \text{ and } B \text{ is nonsingular.} \tag{34}$$

The *congruence class of the regular part B is uniquely determined, so the similarity class of its *cosquare $B^{-*}B$ is also uniquely determined. There are no Type II blocks for A and its Type I blocks are all 1-by-1 if and only if the *cosquare of its regular part is diagonalizable and has only eigenvalues

with unit modulus. Thus, *A is diagonalizable by *congruence if and only if both of the following conditions are satisfied: (a) A and A^* have the same null space, and (b) the *cosquare of the regular part of A is diagonalizable and all its eigenvalues have unit modulus.*

Let A be nonsingular and suppose that its *cosquare $\mathcal{A}$ is diagonalizable and all its eigenvalues have unit modulus. Let S be any nonsingular matrix that diagonalizes $\mathcal{A}$, and consider the forms that (25) and (26) take in this case:

$$A^{-*}A = S\,(C_1 \oplus \cdots \oplus C_q)\,S^{-1}$$

and

$$S^*AS = E_1 \oplus \cdots \oplus E_q,$$

in which $C_j = e^{2i\theta_j}I$ for each $j = 1, ..., q$, $0 \le \theta_1 < \cdots < \theta_q < \pi$, each E_j is the same size as C_j, and each $E_j^{-*}E_j = e^{2i\theta_j}I$. We have $e^{-i\theta_j}E_j = e^{i\theta_j}E_j^* = (e^{-i\theta_j}E_j)^*$, so each matrix $e^{-i\theta_j}E_j$ is Hermitian. Sylvester's Inertia Theorem ensures that $e^{-i\theta_j}E_j$ is *congruent to $I_{n_j^+} \oplus (-I_{n_j^-})$ for nonnegative integers n_j^+ and n_j^- that are determined uniquely by (the *congruence class of) $e^{-i\theta_j}E_j$. It follows that each E_j is *congruent to

$$e^{i\theta_j}I_{n_j^+} \oplus (-e^{i\theta_j}I_{n_j^-}) = e^{i\theta_j}I_{n_j^+} \oplus (e^{i(\theta_j+\pi)}I_{n_j^-}),$$

so A is *congruent to the uniquely determined canonical form

$$\left(e^{i\theta_1}I_{n_1^+} \oplus (e^{i(\theta_1+\pi)}I_{n_1^-})\right) \oplus \cdots \oplus \left(e^{i\theta_q}I_{n_q^+} \oplus (e^{i(\theta_q+\pi)}I_{n_q^-})\right),$$

$0 \le \theta_1 < \cdots < \theta_q < \pi$. The angles θ_j for which $n_j^+ \ge 1$ together with the angles $\theta_j + \pi$ for which $n_j^- \ge 1$ (all $\theta_j \in [0, \pi)$, $j = 1, \ldots, q$) are the canonical angles of order one of A; the corresponding integers n_j^+ and n_j^- are their respective multiplicities.

We have just described how to determine the signs ε_j that occur in (30) when $N = (1, \ldots, 1)$; in this special case, every $|N|$-by-$|N|$ matrix is N-upper Toeplitz, and N-Hermitian matrices are just ordinary Hermitian matrices. *Two square complex matrices of the same size that are diagonalizable by *congruence are *congruent if and only if they have the same canonical angles of order one with the same multiplicities.*

This observation is a special case of a more general fact: over the reals or complexes, each system of forms and linear mappings decomposes uniquely into a direct sum of indecomposables, up to isomorphism of summands [20, Theorem 2]. This special case was rediscovered in [12], which

established uniqueness of the canonical angles and their multiplicities but did not determine them (the signs ε_j remained ambiguous) except in special circumstances, e.g., if the field of values of $e^{i\phi}A$ lies in the open right half plane for some $\phi \in [0, 2\pi)$. Using (34) to introduce generalized inverses and a natural generalized *cosquare, [18] later gave an alternative approach that fully determined the canonical angles and their multiplicities for a complex matrix that is diagonalizable by *congruence.

If A is normal, it is unitarily *congruent to $\Lambda \oplus 0_k$, in which Λ is a diagonal matrix whose diagonal entries are the nonzero eigenvalues of A (including multiplicities). The canonical angles (of order one; no higher orders occur) of a normal matrix are just the principal values of the arguments of its nonzero eigenvalues; the multiplicity of a canonical angle is the number of eigenvalues on the open ray that it determines. Thus, *two normal complex matrices of the same size are *congruent if and only if they have the same number of eigenvalues on each open ray from the origin.* This special case was completely analyzed in [11]. If A is Hermitian, of course, its nonzero eigenvalues are on only two open rays from the origin: the positive half-line and the negative half-line. For Hermitian matrices, the criterion for *congruence of normal matrices is just Sylvester's Inertia Theorem.

Finally, suppose that the *congruence canonical form of a nonsingular n-by-n complex matrix A has only one block; the singular case is analyzed in [9]. If n is odd, that block must be a Type I block $\lambda\Delta_n$ (with $|\lambda| = 1$). If $n = 2m$, it can be either a Type I block or a Type II block $H_{2m}(\mu)$ (with $|\mu| > 1$). Of course, the latter case occurs if and only if the *cosquare of A is similar to $J_m(\mu) \oplus J_m(\bar{\mu}^{-1})$ and $|\mu| > 1$.

Suppose the *cosquare of A is similar to $J_n(\lambda^2)$ with $|\lambda| = 1$. Then the *congruence canonical form of A is $\varepsilon\lambda\Delta_n$ and $\varepsilon \in \{1, -1\}$ can be determined as follows. Let $\hat{A} = \lambda^{-1}A$, let S be such that $\Delta_n^{-1}\Delta_n = S^{-1}(\hat{A}^{-*}\hat{A})S$, and let $M := S^*\hat{A}S\Delta_n^{-1}$. Lemma 5(a) tells us that M is similar to a real matrix. Because $\hat{A}$ is indecomposable under *congruence, in Lemma 5(b) either $k = n$ or $k = 0$:

$$\varepsilon = \begin{cases} -1 & \text{if all the eigenvalues of } M \text{ are negative} \\ 1 & \text{if no eigenvalue of } M \text{ is negative.} \end{cases} \tag{35}$$

Alternatively, we can employ the algorithm described in Section 4. Examine $B := \Delta_n^{-1}S^*\hat{A}S$, which must be nonsingular, upper Toeplitz, and

real, so

$$B = \begin{bmatrix} b_{11} & b_{11}^{(2)} & \dots & b_{11}^{(n)} \\ & b_{11} & \ddots & \vdots \\ & & \ddots & b_{11}^{(2)} \\ 0 & & & b_{11} \end{bmatrix}, \quad b_{11} \neq 0.$$

The reduction described in Lemma 6 is trivial in this case, and it tells us that ε is the sign of b_{11}.

Example 7. Consider

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix},$$

whose *cosquare

$$A^{-*}A = \begin{bmatrix} 1 & 2 \\ -2 & -3 \end{bmatrix}$$

is not diagonal, has -1 as a double eigenvalue, and hence is similar to $J_2(-1)$. Thus, A is *congruent to $\varepsilon i\Delta_2$ with $\varepsilon = \pm 1$. Let $\hat{A} = -iA$ and verify that

$$\begin{bmatrix} 1 & 2i \\ 0 & 1 \end{bmatrix} = \Delta_2^{-*}\Delta_2 = S^{-1}(\hat{A}^{-*}\hat{A})S \quad \text{for } S = \begin{bmatrix} -1 & 0 \\ 1 & i \end{bmatrix}.$$

Then both eigenvalues of $M = S^*\hat{A}S\Delta_2^{-1} = -I_2$ are negative, so (35) ensures that $\varepsilon = -1$ and A is *congruent to $-i\Delta_2$. Alternatively, $B = \Delta_n^{-1}S^*\hat{A}S = -I_2$, so the sign of b_{11} is negative and $\varepsilon = -1$.

Example 8. Consider

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}, \tag{36}$$

whose *cosquare

$$A^{-*}A = \begin{bmatrix} -1 & 2 \\ 0 & -1 \end{bmatrix}$$

is similar to $J_2(-1)$. Thus, A is *congruent to $\varepsilon i\Delta_2$. Let $\hat{A} = -iA$ and verify that

$$\begin{bmatrix} 1 & 2i \\ 0 & 1 \end{bmatrix} = \Delta_2^{-*}\Delta_2 = S^{-1}(\hat{A}^{-*}\hat{A})S \quad \text{for } S = \begin{bmatrix} 1 & 0 \\ 0 & -i \end{bmatrix}.$$

Then both eigenvalues of $M = S^*\hat{A}S\Delta_2^{-1} = -I_2$ are negative, so (35) ensures that $\varepsilon = -1$ and A is *congruent to $-i\Delta_2$.

Example 9. Suppose $|\lambda| = 1$ but $\lambda^2 \neq -1$. Let a denote the real part of λ and consider

$$A = \begin{bmatrix} 0 & \lambda/a \\ \lambda/a & i \end{bmatrix} = \frac{\lambda}{a} \begin{bmatrix} 0 & 1 \\ 1 & a\bar{\lambda}i \end{bmatrix},$$

whose *cosquare

$$A^{-*}A = \lambda^2 \begin{bmatrix} a\lambda i & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & a\bar{\lambda}i \end{bmatrix} = \lambda^2 \begin{bmatrix} 1 & 2a^2 i \\ 0 & 1 \end{bmatrix}$$

is similar to $J_2(\lambda^2)$. Thus, A is *congruent to $\varepsilon\lambda\Delta_2$. Let $\hat{A} = \lambda^{-1}A$ and verify that

$$\begin{bmatrix} 1 & 2i \\ 0 & 1 \end{bmatrix} = \Delta_2^{-*}\Delta_2 = S^{-1}(\hat{A}^{-*}\hat{A})S \quad \text{for } S := \begin{bmatrix} 1 & 0 \\ 0 & 1/a^2 \end{bmatrix}.$$

Then both eigenvalues of

$$M = S^*\hat{A}S\Delta_2^{-1} = \begin{bmatrix} 0 & 1/a^3 \\ 1/a^3 & \bar{\lambda}i/a^4 \end{bmatrix} \begin{bmatrix} -i & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1/a^3 & 0 \\ \star & 1/a^3 \end{bmatrix}$$

have the same sign as a. Thus, (35) ensures that A is *congruent to $\lambda\Delta_2$ if $\operatorname{Re}\lambda > 0$ and to $-\lambda\Delta_2$ if $\operatorname{Re}\lambda < 0$.

6 Canonical pairs

We now explain how to use the canonical matrices in Theorem 1 to obtain the canonical pairs described in Theorem 3.

Proof of Theorem 3(a). Each square matrix A can be expressed uniquely as the sum of a symmetric and a skew-symmetric matrix:

$$A = \mathcal{S}(A) + \mathcal{C}(A), \qquad \mathcal{S}(A) := \tfrac{1}{2}\left(A + A^T\right), \quad \mathcal{C}(A) := \tfrac{1}{2}\left(A - A^T\right). \quad (37)$$

Since $\mathcal{S}(R^TAR) = R^T\mathcal{S}(A)R$ and $\mathcal{C}(R^TAR) = R^T\mathcal{C}(A)R$, any congruence that reduces A to a direct sum

$$R^TAR = B_1 \oplus \cdots \oplus B_k$$

gives a simultaneous congruence of $(\mathcal{S}(A), \mathcal{C}(A))$ that reduces it to a direct sum of pairs

$$(\mathcal{S}(B_1), \mathcal{C}(B_1)) \oplus \cdots \oplus (\mathcal{S}(B_k), \mathcal{C}(B_k)).$$

Theorem 1(a) ensures that A is congruent to a direct sum of blocks of the three types

$$2J_n(0), \quad \Gamma_n, \quad 2H_{2n}(\mu), \tag{38}$$

in which $0 \neq \mu \neq (-1)^{n+1}$ and μ is determined up to replacement by μ^{-1}, and that such a decomposition is unique up to permutation of the direct summands.

Computing the symmetric and skew-symmetric parts of the blocks (38) produces the indicated Type 0, Type I, and Type II canonical pairs in (7).

It remains to prove that the two alternative pairs in (8) may be used instead of the Type II pair

$$\left([J_n(\mu+1) \diagdown J_n(\mu+1)^T],\ [J_n(\mu-1) \diagdown -J_n(\mu-1)^T]\right). \tag{39}$$

First suppose that $\mu = -1$, so n is odd (since $\mu \neq (-1)^{n+1}$) and we have the Type II pair

$$\left([J_n(0) \diagdown J_n(0)^T],\ [J_n(-2) \diagdown -J_n(-2)^T]\right).$$

A simultaneous congruence of this pair via

$$\begin{bmatrix} I_n & 0 \\ 0 & J_n(-2)^{-T} \end{bmatrix}$$

transforms it to the pair

$$\left(\begin{bmatrix} 0 & J_n(0)^T J_n(-2)^{-T} \\ J_n(-2)^{-1} J_n(0) & 0 \end{bmatrix},\ \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}\right). \tag{40}$$

Since $J_n(-2)^{-1}J_n(0)$ is similar to $J_n(0)$, there is a nonsingular matrix S such that $S^{-1}J_n(0)S = J_n(-2)^{-1}J_n(0)$. Then a simultaneous congruence of (40) via

$$\begin{bmatrix} S^{-1} & 0 \\ 0 & S^T \end{bmatrix}$$

transforms it to the second of the two alternative pairs in (8).

Now suppose that $\mu \neq -1$. Let S be a nonsingular matrix such that

$$S^{-1}J_n(\mu-1)J_n(\mu+1)^{-1}S = J_n(\nu), \quad \nu := \frac{\mu-1}{\mu+1}. \tag{41}$$

A simultaneous congruence of the pair (39) via

$$\begin{bmatrix} J_n(\mu+1)^{-1}S & 0 \\ 0 & S^{-T} \end{bmatrix}$$

transforms it to the first alternative pair in (8). The definition (41) ensures that $\nu \neq 1$; $\nu \neq -1$ since $\mu \neq 0$; and $\nu \neq 0$ if n is odd since $\mu \neq (-1)^{n+1}$. Because μ is determined up to replacement by μ^{-1}, ν is determined up to replacement by

$$\frac{\mu^{-1}-1}{\mu^{-1}+1} = \frac{1-\mu}{1+\mu} = -\nu.$$

$\square$

Proof of Theorem 3(b). Each square complex matrix has a *Cartesian decomposition*

$$A = \mathcal{H}(A)+i\mathcal{K}(A), \qquad \mathcal{H}(A) := \tfrac{1}{2}\left(A + A^*\right), \quad \mathcal{K}(A) := \tfrac{i}{2}\left(-A + A^*\right) \tag{42}$$

in which both $\mathcal{H}(A)$ and $\mathcal{K}(A)$ are Hermitian. Moreover, if $\mathcal{H}'$ and $\mathcal{K}'$ are Hermitian matrices such that $A = \mathcal{H}'+i\mathcal{K}'$, then $\mathcal{H}' = \mathcal{H}(A)$ and $\mathcal{K}' = \mathcal{K}(A)$. Since $\mathcal{H}(R^*AR) = R^*\mathcal{H}(A)R$ and $\mathcal{K}(R^*AR) = R^*\mathcal{K}(A)R$, any *congruence that reduces A to a direct sum

$$R^*AR = B_1 \oplus \cdots \oplus B_k$$

gives a simultaneous *congruence of $(\mathcal{H}(A), \mathcal{K}(A))$ that reduces it to a direct sum of pairs

$$(\mathcal{H}(B_1), \mathcal{K}(B_1)) \oplus \cdots \oplus (\mathcal{H}(B_k), \mathcal{K}(B_k)).$$

Theorem 1(b) ensures that A is congruent to a direct sum of blocks of the three types

$$2J_n(0), \ \lambda\Delta_n, \ \text{and } H_{2n}(\mu), \ \text{in which } |\lambda| = 1 \text{ and } |\mu| > 1 \tag{43}$$

and that such a decomposition is unique up to permutation of the direct summands. The Cartesian decomposition of $2J_n(0)$ produces the Type 0 pair in (9).

Consider the Type I block $\lambda\Delta_n$ with $|\lambda| = 1$. If a matrix F_n is nonsingular and $F_n^{-*}F_n$ is similar to $J_n(\lambda^2)$, then $\lambda\Delta_n$ is *congruent to $\pm F_n$. Suppose $\lambda^2 \neq -1$. Then

$$c := i\frac{1-\lambda^2}{1+\lambda^2} = i\frac{\bar{\lambda}(1-\lambda^2)}{\bar{\lambda}(1+\lambda^2)} = i\frac{\bar{\lambda}-\lambda}{\bar{\lambda}+\lambda} = \frac{\operatorname{Im}\lambda}{\operatorname{Re}\lambda}$$

is real and

$$\lambda^2 = \frac{1+ic}{1-ic}.$$

Consider the symmetric matrix

$$F_n := \Delta_n(1+ic, i) = P_n \left(I_n + iJ_n(c)\right). \tag{44}$$

Then

$$F_n^{-*} = \overline{F_n^{-1}} = \left(I_n - iJ_n(c)\right)^{-1} P_n$$

and

$$\begin{aligned} F_n^{-*}F_n &= \left(I_n - iJ_n(c)\right)^{-1} P_n P_n \left(I_n + iJ_n(c)\right) \\ &= \left(\left(1-ic\right) I_n - iJ_n(0)\right)^{-1} \left(\left(1+ic\right) I_n + iJ_n(0)\right) \\ &= \lambda^2 \left(I_n + a_1 J_n(0) + a_2 J_n(0)^2 + a_3 J_n(0)^3 + \cdots\right), \end{aligned}$$

in which $a_1 = 2i\left(1+c^2\right)^{-1} \neq 0$. Thus, $F_n^{-*}F_n$ is similar to $J_n(\lambda^2)$, so $\lambda\Delta_n$ is *congruent to $\pm\Delta_n(1+ic, i)$. The Cartesian decomposition of $\pm\Delta_n(1+ic, i)$ produces the first Type I pair in (9).

Now suppose that $\lambda^2 = -1$ and consider the real matrix

$$G_n := \begin{cases} \begin{bmatrix} 0 & & & & & 1 \\ & & & & \iddots & 1 \\ & & & 1 & \iddots & \\ & & -1 & 1 & & \\ & \iddots & \iddots & & & \\ -1 & 1 & & & & 0 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} m \\ \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} m \end{matrix} & \text{if } n = 2m, \\[2ex] \begin{bmatrix} 0 & & & & & & 1 \\ & & & & & \iddots & 1 \\ & & & & 1 & \iddots & \\ & & & \boxed{1} & 1 & & \\ & & 1 & 0 & & & \\ & \iddots & \iddots & & & & \\ 1 & 0 & & & & & 0 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} m \\ \}\, 1 \\ \left.\vphantom{\begin{matrix}1\\1\\1\end{matrix}}\right\} m \end{matrix} & \text{if } n = 2m+1 \end{cases} \tag{45}$$

(the boxed unit is at the center). Since $G_n^{-T}G_n$ is similar to $J_n(-1)$, $\lambda\Delta_n$ is *congruent to $\pm G_n$. The Cartesian decomposition of $\pm G_n$ produces the second Type I pair in (9).

The Type II block $H_{2n}(\mu)$ with $|\mu| > 1$ is *congruent to $H_{2n}(\bar{\mu}^{-1})$ because their *cosquares are both similar to $J_n(\mu) \oplus J_n(\bar{\mu}^{-1})$. Represent $\bar{\mu}^{-1}$,

a point in the open unit disk with the origin is omitted, as

$$\bar{\mu}^{-1} = \frac{1+i\nu}{1-i\nu},$$

in which ν is in the open upper half plane with the point i omitted. In fact,

$$\nu = \frac{2\,\mathrm{Im}\,\mu + i\left(|\mu|^2 - 1\right)}{|\mu+1|^2} := a + ib \neq i, \quad b > 0, \quad a, b \in \mathbb{R}. \tag{46}$$

We have

$$\begin{aligned}(I_n + iJ_n(\nu))\,(I_n - iJ_n(\nu))^{-1} &= ((1+i\nu)\,I_n + iJ_n(0))\,((1-i\nu)\,I_n - iJ_n(0))^{-1} \\ &= \bar{\mu}^{-1}\left(I_n + a_1 J_n(0) + a_2 J_n(0)^2 + a_3 J_n(0)^3 + \cdots\right),\end{aligned}$$

in which $a_1 = 2i\left(1+\nu^2\right)^{-1} \neq 0$. Thus,

$$(I_n + iJ_n(\nu))\,(I_n - iJ_n(\nu))^{-1}$$

is similar to $J_n(\bar{\mu}^{-1})$. Let S be a nonsingular matrix such that

$$S^{-1} J_n(\bar{\mu}^{-1}) S = (I_n + iJ_n(\nu))\,(I_n - iJ_n(\nu))^{-1}$$

and compute the following *congruence of $H_{2n}(\bar{\mu}^{-1})$:

$$\begin{aligned}&\begin{bmatrix} S\,(I_n - iJ_n(\nu)) & 0 \\ 0 & S^{-*} \end{bmatrix}^* \begin{bmatrix} 0 & I_n \\ J_n(\bar{\mu}^{-1}) & 0 \end{bmatrix} \begin{bmatrix} S\,(I_n - iJ_n(\nu)) & 0 \\ 0 & S^{-*} \end{bmatrix} \\ &\qquad = \begin{bmatrix} 0 & I_n + iJ_n(\nu)^* \\ I_n + iJ_n(\nu) & 0 \end{bmatrix}.\end{aligned}$$

The Cartesian decomposition of this matrix produces the Type II pair in (9) with the parameters a and b defined in (46). □

For a given λ with $|\lambda| = 1$, the $\pm$ signs associated with the Type I canonical pairs in (9) can be determined using the algorithms in either Section 3 or Section 4.

For example, suppose $n = 2$ and $\lambda = i$. The matrix

$$G_2 = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}$$

defined in (45) was analyzed in Example 8. We found that $i\Delta_2$ is *congruent to $-G_2$, so the *congruence canonical pair associated with $i\Delta_2$ is

$$-\left(\Delta_2(0,1), \Delta_2(1,0)\right) = -\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right).$$

As a second example, suppose $n = 2$ and $|\lambda| = 1$, but $\lambda^2 \neq -1$. Let $\lambda = a + ib$ $(a, b \in \mathbb{R})$. The matrix

$$F_2 = \Delta_2(1 + ib/a, i) = \Delta_2(\lambda/a, i) = \begin{bmatrix} 0 & \lambda/a \\ \lambda/a & i \end{bmatrix}$$

defined in (44), was analyzed in Example 9. We found that $\lambda\Delta_2$ is *congruent to F_2 if $a > 0$, and to $-F_2$ if $a < 0$. Thus, the *congruence canonical pair associated with $\lambda\Delta_2$ is

$$\left(\Delta_2(1,0),\ \Delta_2\left(b/a, 1\right)\right) = \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & b/a \\ b/a & 1 \end{bmatrix}\right)$$

if $\operatorname{Re}\lambda > 0$, and is

$$-\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & b/a \\ b/a & 1 \end{bmatrix}\right)$$

if $\operatorname{Re}\lambda < 0$.

References

[1] C. S. Ballantine, Cosquares: complex and otherwise, *Linear and Multilinear Algebra* 6 (1978) 201–217.

[2] B. Corbas and G. D. Williams, Bilinear forms over an algebraically closed field, *J. Pure Appl. Algebra* 165 (3) (2001) 225–266.

[3] P. Gabriel, Appendix: degenerate bilinear forms, *J. Algebra* 31 (1974) 67–72.

[4] F. R. Gantmacher, *The Theory of Matrices*, vol. I , Chelsea, New York, 2000.

[5] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.

[6] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, New York, 1991.

[7] R. A. Horn and G. Piepmeyer, Two applications of the theory of primary matrix functions, *Linear Algebra Appl.* 361 (2003) 99-106.

[8] R. A. Horn and V. V. Sergeichuk, Congruence of a square matrix and its transpose, *Linear Algebra Appl.* 389 (2004) 347–353.

[9] R. A. Horn and V. V. Sergeichuk, A regularizing algorithm for matrices of bilinear and sesquilinear forms, *Linear Algebra Appl.* 412 (2006) 380–395.

[10] L. K. Hua, On the theory of automorphic functions of a matrix variable II—The classification of hypercircles under the symplectic group, *Amer. J. Math.* 66 (1944) 531-563.

[11] Kh. D. Ikramov, On the inertia law for normal matrices, *Doklady Math.* 64 (2001) 141–142.

[12] C. R. Johnson and S. Furtado, A generalization of Sylvester's law of inertia, *Linear Algebra Appl.* 338 (2001) 287–290.

[13] P. Lancaster, L. Rodman, Canonical forms for symmetric/skew-symmetric real matrix pairs under strict equivalence and congruence, *Linear Algebra Appl.* 406 (2005) 1–76.

[14] P. Lancaster and L. Rodman, Canonical forms for Hermitian matrix pairs under strict equivalence and congruence, *SIAM Review* 47 (2005) 407-443.

[15] J. M. Lee and D. A. Weinberg, A note on canonical forms for matrix congruence, *Linear Algebra Appl.* 249 (1996) 207–215.

[16] C. Riehm, The equivalence of bilinear forms, *J. Algebra* 31 (1974) 45–66.

[17] C. Riehm and M. Shrader–Frechette, The equivalence of sesquilinear forms, *J. Algebra* 42 (1976) 495–530.

[18] D. W. Robinson, An alternative approach to unitoidness, *Linear Algebra Appl.* 413 (2006) 72-80.

[19] A. V. Roiter, Bocses with involution, in: *Representations and Quadratic Forms* (Ju. A. Mitropol′skii, Ed.), Inst. Mat. Akad. Nauk Ukrain. SSR, Kiev, 1979, 124–128 (in Russian).

[20] V. V. Sergeichuk, Classification problems for system of forms and linear mappings, *Math. USSR, Izvestiya* 31 (3) (1988) 481–501.

[21] R. C. Thompson, Pencils of complex and real symmetric and skew matrices, *Linear Algebra Appl.* 147 (1991) 323–371.

[22] G. E. Wall, On the conjugacy classes in the unitary, symplectic and orthogonal groups, *J. Aust. Math. Soc.* 3 (1963) 1-62.

[23] W. C. Waterhouse, The number of congruence classes in $M_n(F_q)$, *Finite Fields Appl.* 1 (1995) 57–63.